# Claims

[c1]  1. A system for providing information related to one or more probe sets, the system comprising:

an input manager constructed and arranged to receive from a user a selection of one or more probe-set identifiers associated with one or more probe sets, wherein the probe-set identifiers include one or more annotation terms;

a determiner constructed and arranged to correlate the user-selected probe-set identifiers with one or more biological sequences; and

a correlator constructed and arranged to correlate the one or more biological sequences with a first set of probe sets.

[c2]  2. The system of claim 1 , further comprising:

an output manager constructed and arranged to identify the first set of probe sets to the user.

[c3]  3. The system of claim 2 , wherein:

the output manager identifies the first set of probe sets in a tabular format.

[c4]  4. The system of claim 1 , wherein:

the input manager receives the user selection over the Internet.

[c5]  5. The system of claim 1 , wherein:

the one or more probe sets include probes of a synthesized probe array.

[c6]  6. The system of claim 1 , wherein:

the one or more probe sets include probes of a spotted probe array.

[c7]  7. The system of claim 1 , wherein:

the probe sets are capable of the identification of at least one biological molecule.

[c8]  8. The system of claim 1 , wherein:

the biological sequences include one or more of a whole or a part of a consensus, EST, gene, or SIF sequence.

[c9]

9. The system of claim 1 , wherein:

the annotation terms are descriptive of any one or more of molecular function, cellular location, tissue type, or biological process.

[c10]     10. A method for providing information related to one or more probe sets, the method comprising:

the act of receiving from a user a selection of one or more probe-set identifiers associated with one or more probe sets, wherein the probe-set identifiers include one or more annotation terms;

the act of correlating the user-selected probe-set identifiers with one or more biological sequences; and

the act of correlating the one or more biological sequences with a first set of probe sets.

[c11]     11. The method of claim 10, further comprising:

the act of identifying the first set of probe sets to the user.

[c12]     12. The method of claim 10, wherein:

the probe sets are capable of the identification of at least one biological molecule.

[c13]     13. The method of claim 10, wherein:

the biological sequences include one or more of a whole or a part of a consensus, EST, gene, or SIF sequence.

[c14]     14. The method of claim 10, wherein:

the one or more probe sets include probes of a synthesized probe array.

[c15]     15. A genomic portal system for providing information related to one or more probe sets, the system comprising:

an application server comprising an input manager constructed and arranged to receive from a user a selection of one or more probe-set identifiers associated with one or more probe sets, wherein the probe-set identifiers include one or more annotation terms, a determiner constructed and arranged to correlate the user-selected probe-set identifiers with one or more biological sequences, and a correlator constructed and arranged to correlate the one or more biological sequences with a first set of probe sets; and

a network server comprising an output manager constructed and arranged to identify the first set of probe sets to the user.

[c16]    16. The system of claim 15 , wherein:

the network server further comprises an input manager constructed and arranged to receive user input; and

the system further comprises one or more user computers constructed and arranged to enable a user to provide the user selection of one or more probe-set identifiers to the network server.

[c17]    17. The system of claim 15 , wherein:

the annotation terms are descriptive of any one or more of molecular function, cellular location, tissue type, or biological process.

[c18]    18. The system of claim 15 , wherein:

the output manager identify the first set of probe sets to the user via the Internet.

[c19]    19. A method for providing information related to one or more probe sets, wherein each probe-set has one or more identifiers comprising:

the act of receiving from a user a selection of a first set of one or more of the probe-set identifiers, wherein each probe-set is capable of the identification of a biological molecule;

the act of correlating the first set of probe-set identifiers with a first set of one or more data;

the act of correlating each of the first set of data with a second set of one or more data; and

the act of providing the second set of data to the user.

[c20]

20. A system for providing information related to one or more probe sets, the system comprising:

an input manager constructed and arranged to receive from a user a selection of one or more probe-set identifiers associated with one or more probe sets, wherein the probe-set identifiers include one or more biological sequences;

a database manager constructed and arranged to periodically update one or

more local genomic databases; and

a data generator constructed and arranged to cluster the one or more biological
sequences based at least in part on at least one of the local genomic databases.

[c21]     21. The system of claim 20 , further comprising:

a data processor constructed and arranged to format the clustered data for use
in a graphical user interface.

[c22]     22. The system of claim 21 , wherein:

the graphical user interface includes one or more graphical elements
representing protein family data, sequence alignment, or both.

[c23]     23. The system of claim 22 , wherein:

the sequence alignment graphical element includes a representation of a
consensus sequence specific to a protein family.

[c24]     24. The system of claim 21 , wherein:

the format includes XML or HTML format.

[c25]     25. The system of claim 20 , further comprising:

a data processor constructed and arranged to format the clustered data for
storage.

[c26]     26. The system of claim 25 , further comprising:

an output manager constructed and arranged to accept the formatted data from
the data processor and provide the data to a user for storage.

[c27]     27. The system of claim 21 , further comprising:

an output manager constructed and arranged to accept the formatted data from
the data processor and provide the data to a user for display.

[c28]     28. The system of claim 20 , wherein:

the one or more biological sequences comprises a protein sequence.

[c29]

29. The system of claim 20 , wherein:

the probe-set identifiers include one or more nucleotide sequences;

the system further comprises a determiner constructed and arranged to

correlate the one or more nucleotide sequences with one or more protein sequences; and

the data generator clusters the one or more protein sequences based at least in part on at least one of the local genomic databases.

[c30]    30. The system of claim 29 , wherein:

the nucleotide sequence includes one or more of a gene, an EST, a consensus, or an SIF sequence.

[c31]    31. The system of claim 20 , wherein:

the data generator clusters the one or more protein sequences based, at least in part, on a learning algorithm.

[c32]    32. The system of claim 31 , wherein:

the learning algorithm comprises a Hidden Markov Model.

[c33]    33. The system of claim 31 , wherein:

the learning algorithm uses a training dataset that includes one or more protein sequences grouped by family.

[c34]    34. A method for providing information related to one or more probe sets, the method comprising:

the act of receiving from a user a selection of one or more probe-set identifiers associated with one or more probe sets, wherein the probe-set identifiers include one or more biological sequences;

the act of periodically updating one or more local genomic databases; and

the act of clustering the one or more biological sequences based at least in part on at least one of the local genomic databases.

[c35]    35. The method of claim 34 , further comprising:

the act of formatting the clustered data for use in a graphical user interface.

[c36]    36. The method of claim 35 , wherein:

the graphical user interface includes one or more graphical elements representing protein family data, sequence alignment, or both.

[c37]    37. The method of claim 34 , wherein:

the one or more biological sequences comprises a protein sequence.

[c38] 38. The method of claim 34 , wherein:

the probe-set identifiers include one or more nucleotide sequences; and

the method further comprises the acts of correlating the one or more nucleotide sequences with one or more protein sequences, and clustering the one or more protein sequences based at least in part on at least one of the local genomic databases.

[c39] 39. A genomic portal system for providing information related to one or more probe sets, the system comprising:

(1) an application server comprising (a) an input manager constructed and arranged to receive from a user a selection of one or more probe-set identifiers associated with one or more probe sets, wherein the probe-set identifiers include one or more biological sequences, (b) a database manager constructed and arranged to periodically update one or more local genomic databases, (c) a data generator constructed and arranged to cluster the one or more biological sequences based at least in part on at least one of the local genomic databases, and (d) a data processor constructed and arranged to format the clustered data; and

(2) a network server comprising an output manager constructed and arranged to accept the formatted cluster data from the data processor and provide the data to a user.

[c40] 40. A genomic portal system, comprising:

(1) an application server constructed and arranged to (a) receive from a user a plurality of probe-set identifiers associated with a plurality of probe sets, wherein the probe-set identifiers are included in a batch file, (b) periodically update one or more local genomic databases, and (c) correlate each of the probe-set identifiers with data based on at least one of the local genomic databases; and

(2) a network server constructed and arranged to provide the data to the user.

[c41] 41. The system of claim 40 , wherein:

the probe set identifiers include any one or more of accession number,

manufacturer-defined probe set identifier, biological sequence, or annotation term.

[c42]    42. The system of claim 40 , wherein:
each probe set includes one or more probes capable of identifying a biological molecule.

[c43]    43. The system of claim 42 , wherein:
the probes are included in a synthesized probe array.

[c44]    44. The system of claim 42 , wherein:
the probes are included in a spotted probe array.

[c45]    45. The system of claim 40 , wherein:
the application server further is constructed and arranged to parse the batch file into sets of data corresponding to each probe set identifier.

[c46]    46. The system of claim 45 , wherein:
the application server further is constructed and arranged to perform user-specified operations on the sets of data.

[c47]    47. The system of claim 46 , wherein:
the user-specified operations include any one or more of obtaining probe set records; sorting probe sets; performing similarity searches; processing array content queries; or obtaining protein domain, sequence homology, complex membership, pathway information, biological system role, or interaction with other proteins or biological molecules.

[c48]    48. A method, comprising:
the act of receiving from a user a plurality of probe-set identifiers associated with a plurality of probe sets, wherein the probe-set identifiers are included in a batch file;
the act of periodically updating one or more local genomic databases;
the act of correlating each of the probe-set identifiers with data based on at least one of the local genomic databases; and
the act of providing the data to the user over a network.

[c49]    49. A system for providing information related to one or more probe sets, comprising:

an input manager constructed and arranged to receive one or more probe-set identifiers associated with one or more probe sets;

a determiner constructed and arranged to correlate the probe-set identifiers with a first set of data; and

a correlator constructed and arranged to correlate the first set of data with a second set of data;

wherein the probe-set identifiers include one or more of accession number, manufacturer-defined probe set identifier, biological sequence, or annotation term, and further wherein the first set of data includes biological sequence or structure data, and still further wherein the second set of data includes a set identifier or protein information.

[c50]    50. The system of claim 49 , wherein:

the protein information includes biological process, molecular function, or cellular location information.

[c51]    51. The system of claim 50 , wherein:

the protein information includes information related to protein domain, sequence homology, complex membership, pathway, biological system role, or interaction with other proteins or biological molecules.

[c52]    52. A system for providing information related to one or more probe sets, comprising:

an input manager constructed and arranged to receive one or more probe-set identifiers associated with one or more probe sets; and

a manager constructed and arranged to correlate the probe-set identifiers with a set of data;

wherein the probe-set identifiers include one or more of accession number, manufacturer-defined probe set identifier, biological sequence, or annotation term, and further wherein the set of data includes protein information.

[c53]    53. The system of claim 52 , wherein:

the protein information includes biological process, molecular function, or

cellular location information.

[c54]    54. The system of claim 52, wherein:

the protein information includes information related to protein domain,

sequence homology, complex membership, pathway, biological system role, or

interaction with other proteins or biological molecules.

[c55]    55. A method, comprising:

the act of receiving one or more probe-set identifiers associated with one or

more probe sets; and

the act of correlating the probe-set identifiers with a set of data;

wherein the probe-set identifiers include one or more of accession number,

manufacturer-defined probe set identifier, biological sequence, or annotation

term, and further wherein the set of data includes protein information.

[c56]    56. The method of claim 55, wherein:

the protein information includes biological process, molecular function, or

cellular location information.

[c57]    57. The method of claim 55, wherein:

the protein information includes information related to protein domain,

sequence homology, complex membership, pathway, biological system role, or

interaction with other proteins or biological molecules.

[c58]    58. A genomic portal system, comprising:

an application server comprising an input manager constructed and arranged to

receive one or more probe-set identifiers associated with one or more probe

sets, and a manager constructed and arranged to correlate the probe-set

identifiers with a set of data; and

a network server constructed and arranged to provide the data to the user;

wherein the probe-set identifiers include one or more of accession number,

manufacturer-defined probe set identifier, biological sequence, or annotation

term, and further wherein the set of data includes protein information.

[c59]    59. The method of claim 58, wherein:

the protein information includes biological process, molecular function, or

cellular location information.

[c60]     60. The method of claim 58 , wherein:

the protein information includes information related to protein domain,

sequence homology, complex membership, pathway, biological system role, or

interaction with other proteins or biological molecules.